

SEQUENCE LISTING - WEE virus strain 71V-1658 - SEQ ID NO 1

LOCUS AF214040 11484 bp RNA VRL 27-JAN-2000  
 DEFINITION Western equine encephalomyelitis virus strain 71V-1658  
 nonstructural polyprotein and structural polyprotein genes,  
 complete cds.  
 ACCESSION AF214040 AF143811  
 VERSION AF214040.1 GI:6760410  
 KEYWORDS .  
 SOURCE Western equine encephalomyelitis virus.  
 ORGANISM Western equine encephalomyelitis virus Viruses; ssRNA  
 positive-strand viruses, no DNA stage; Togaviridae; Alphavirus.  
 REFERENCE 1 (bases 1 to 11484)  
 AUTHORS Netolitzky,D.J., Schmaltz,F.L., Parker,M.D., Rayner,G.A.,  
 Fisher,G.R., Trent,D.W., Bader,D.E. and Nagata,L.P.  
 TITLE Complete genomic RNA sequence of western equine  
 encephalitis virus and expression of the structural genes  
 JOURNAL J. Gen. Virol. 81 Pt 1, 151-159 (2000)  
 MEDLINE 20109043  
 REFERENCE 2 (bases 1 to 11484)  
 AUTHORS Netolitzky,D.J., Schmaltz,F.L., Parker,M.D., Rayner,G.A.,  
 Fisher,G.R., Trent,D.W., Bader,D.E. and Nagata,L.P.  
 TITLE Direct Submission  
 JOURNAL Submitted (08-DEC-1999) Medical Countermeasures Section,  
 Defence Research Establishment Suffield, P.O. Box 4000, Stn Main,  
 Medicine Hat, Alberta T1A 8K6, Canada  
 FEATURES Location/Qualifiers  
 source 1..11484  
 /organism="Western equine encephalomyelitis virus"  
 /strain="71V-1658"  
 /db\_xref="taxon:11039"  
 5'UTR <1..24  
 misc\_feature 1..11484  
 /note="sequence of this entry added to the sequence of  
 AF214039, strain CBA87, constitutes the Western equine  
 encephalomyelitis virus complete genome"  
  
 CDS 25..7428  
 /note="encodes nsP1, nsP2, nsP3 and nsP4 proteins"  
 /codon\_start=1  
 /transl\_except=(pos:5581..5583,aa:OTHER)  
 /product="nonstructural polyprotein"  
 /protein\_id="AAF28339.1"  
 /db\_xref="GI:6760411"  
  
 /translation="MERIHVDLDADSPYVKSLQRTFPQFEIARQVTDNDHANARAFS  
 HVATKLIIESEVDRDQVILDIGSAPVRHAHSNHRYHCICPMISAEDPDRLQRYAERLKK

10023649-122101

10023649-122101

SDITDKNIASKAADLLEVMSTPDAETPSLCMHTDATCRYFGSVAVYQDVYAVHAPTSI  
YHQALKGVRTIYWIGFDTPPFMYKNMAGSYPTYNTNWADERVLEARNIGLGNSDLQES  
RLGKLSILRKKRLQPTNKIIIFSVGSTIYTEDRSLLRSWHLPNVFHLKGKSNFTGRCGT  
IVSCEGYVIKKITISPGLYGKVENLASTMHREGFLSCKVTDTLRGERVSFAVCTYVPA  
TLCDQMTGILATDVSVDDAQKLLVGLNQRIVVNGRTQRNTNTMQNYLLPVVAQAFSRW  
AREHRADLDDEKELGVRERTLTMGCCWAFKTQKITSYKKPGTQTIKKVPAVFDSFVI  
PRLTSHGLDMGFRRRLKLLLEPTVKPAPAITMADVEHLRGLQQEAEVAAAEEIREAL  
PPLLPEIEKETVEAEVDLIMQEAGAGSVETPRGHIRVTSYPGEEKIGSYAILSPQAVL  
NSEKLACIHPLAEQVLVMTHKGRAGRYKVEPYHGKVIVPEGTAVPVQDFQALSESATI  
VFNEREFVNRYLHHIAINGGALNTDEEYYKTVKTQDSEYVFDIDARKCVKREDAGP  
LCLTGDLVDPPFHEFAYESLKTRPAAPHKVPTIGVYGVPGSGKSGI I KSAVTKKDLVV  
SAKKENCAEIIIRDVRRMRMDVAARTVDSVLLNGVKHPVNTLYIDEAFACHAGTLLAL  
IAIVKPKKVVL CGDPKQCGFFNMCLKVHFNHDICTEVYHKSISRCTQTVTAIVSTL  
FYDKRMKTVNPCADKIIIDTTGTTKPHKDDLILTCFRGWVKQLQIDYKNHEIMTAAAS  
QGLTRKGVYAVRYKVNENPLYSQTSEHVNVLTRTEKRIWKTLAGDPWIKTLTAKYP  
GDFTASLDDWQREHDAIMARVLDKPQTADVFNKVNVCWAKALEPVLATANIVLTRQQ  
WETLHPFKHDRAYSPEMALNFFCTRFFGVLDLDSGLFSAPTVALTYRDQHWDNSPGKNM  
YGLNREVAKELSRRYPCITKAVD TGRVADIRNNTIKDYSPTINVVPLNRRLPHSLIVD  
HKGQGTDDHSGFLSKMKGKSVLVIGDPISIPGKKVESMGPLPTNTIRCDL DLGIPSHV  
GKYDII FVNVRTPYRNHHYQQCEDHAIHHSMLTCKAVHHLNTGGTCVAIGYGLADRAT  
ENIITAVARSFRFTRVCQPKNTAENTEVLFVFFGKDNGNHTHDQDRLGVVLDNIYQGS  
TRYEAGRAPAYRVIRGDISKSADQAI VNAANSKGQPGSGVCGALYRKWPAAFDROPIA  
VGTARLVKHEPLIIHAVGPNFSKMPEPEGDLKLAAAYMSIASIVNAERITKISVPLLS  
TGIYSGGKDRVMQSLHHLFTAFTDADVTIYCLDKQWETRIIEAIHRKESVEILDDD  
KPV DIDLVRVHPNSSLAGRPGYSVNEGKLYSYLEGTRFHQTAKDIAEIHAMWPNKSEA

10023649.122101

NEQICLYILGESMSSIRSKCPVEESEASAPPHTLPCLCNYAMTAERVYRLRSAKKEQF  
AVCSSFLLPKYRITGVQKLQCSKPVLFSGVVPPAVHPRKYAEIILETPPPPATTTVIC  
EPTVPERIPSPVISRAPSAESLLSLGGVSFSSSATRSSTAWSDYDRRFVVTADVHQAN  
TSTWSIPSAPGLDVQLPSDVTDSHWSIPSASGFVVRTPSVQDLTAECAKPRGLAEIMQ  
DFNTAPFQFLSDYRPVPAPRRRPIPSPRSTASAPPVPKPRRTKYQQPPGVARAISEAE  
LDEYIRQHSNXRYEAGAYIFSSETGQGHLQQKSVRQCKLQEPILDRAVHEKYYAPRLD  
LEREKMLQKKLQLCASEGNRSRYQSRKVENMKAITAERLISGLGTYLSSEVNPVECYR  
VNYPVPIYSSTVINRFTSAEVAVKTCNLVIQENYPTVASYCITDEYDAYLDMVDGASC  
CLDTATFCPAKLRSYPKKHSYLQPEIRSAVPSPIQNTLQNVLAATKRNCNVTQMREL  
PVLDSAAFNVDCFKKYACNDEYWDTFRDNPIRLTENVTOYVTKLKGPKAAALFANTH  
NLKPLQEIPMDQFVMDLKRDKVTPGTHTEERPKVQVIQAADPLATAYLCGIHREL  
RRLNAVLLPNIHTLFDMSAEDFDAIIAEHFHHGDPVLETDIASFDKSEDDAIAISALM  
ILEDLGVDQPLLDLIEAAFGNITSVHLPTGTRFKFGAMMKSGMFLTLFVNTLVNIMIA  
SRVLRERLTTSACAASIGDDNIVHGVVSDTLMAERCATWLNMEVKIIDAVIGIKAPYF  
CGGFILVDQITGTACRVADPLKRLFKLGKPLPVDDTQDCDRRRALHDEAMRWNRRIGIT  
DELVKAVESRYEIIILAGLIITSLSTLAESVKNFKSIRGSPITLYG"

CDS 7473..11183  
/note="encodes nucleocapsid, E3, E2, 6K and E1 proteins"  
/codon\_start=1  
/product="structural polyprotein"  
/protein\_id="AAF28340.1"  
/db\_xref="GI:6760412"

/translation="MFPYPQLNFPPVYPTNPMAYRDPNPPRRRWRPFRPPLAAQIEDL  
RRSIVNLTFKQRSPNPPPGPPPKKKKSAPKPKPTQPKKKKQQAARTKRKPKPGKRQRM  
CMKLESDKTFPIMLNGQVNGYACVVGGRLMKPLHVEGKIDNEQLAAVKLKKASMYDLE  
YGDVPQNMKSDTLQYTSKPPGFYNWHHGAQYENGRFTVPRGVGGKGDSEGRPILDNR  
GRVVAIVLGGANEGTRTALSVVVTWNQKGVITRDTPEGSEPWSLVTALCVLSNVTFFPCD

10023649.12201

KPPVCYSLTPERTLDVLEENVDPNPYDTLLENVLKCPSRRPKRSITDDFTLTSPYLGF  
CPYCRHSTPCFSPIKIENVWDESDDGSIRIQVSAQFGYNQAGTADVTKFRYMSFDHHD  
DIKEDSMEKIAISTSGPCRRLGHKGYFLLAQCPPGDSVTVSITSGASENSCTVEKKIR  
RKFBVGREEYLFPPVHGKLVKCHVYDHLKETSAGYITMHRPGPHAYKSYLEEASGEVYI  
KPPSGKNVTYECKCGDYSTGIVSTRTKMNGCTKAKQCIAYKSDQTKWVFNSPDLIRHT  
DHSVQGKLHIPFRLTPTVCPVPLAHTPTVTKWFKGITLHLTAMRPTLLTTRKLGLRAD  
ATAEWITGSTSRNFSVGREGLEYVWGNHEPVRVWAQESAPGDPHGWPHHEIIHYHRH  
PVYTVIVLCGVALAILVGTASSAACIAKARRDCLTPYALAPNATVPTALAVLCCIRPT  
NAETFGETLNHLWFNNQPFLWAQLCIPLAALVILFRFCFSCCMPFLLVAGVCLGKVDAF  
EHATTVPNPVPGIPYKALVERAGYAPLNLEITVVSSELTPTSTNKEYVTCKFHTVIPSPQ  
VKCCGSLECKASSKADYTCRVFGGVYPFMWGGAQCFCDSSENTQLSEAYVEFAPDCTID  
HAVALKVHTAALKVGLRIVYGNTTAHLDTFVNGVTPGSSRDLKVIAGPISAAFSFPDH  
KVVIRKGLVYNYDFPEYGAMKPGAFGDIQASSLDATDIVARTDIRLLKPSVKNIHVPY  
TQAVSGYEMWKNNNGRPLQETAPFGCKIEVEPLRASNCAYGHIPIISIDIPDAAFVRSS  
ESPTILEVSCTVADCIYSADFGGSLTLQYKADREGHCPVHSHSTTAVLKEATTHVTAV  
GSITLHFSTSSPQANFIVSLCGKKSTCNAECKPPADHIIGEPHKVDQEFQAAVSKTSW  
NWLLALFGGASSLIVVGLIVLVCSSMLINTRR"

3'UTR

11184..11484

BASE COUNT      3251 a      2842 c      2808 g      2583 t  
ORIGIN

1	accctacaaa	ctaatacgatc	caatatggaa	agaattcacg	ttgacttaga	tgctgacagc
61	ccgtatgtca	agtcggttaca	gcggacggtt	ccacaatttg	agatcgaagc	aaggcaggtc
121	actgacaatg	accatgccaa	tgccagagcg	ttttcgcatg	tggaacaaaa	gctcattgag
181	agcgaagtcg	accgggacca	agttatcttg	gacattggaa	gtgcgcccgt	cagacatgca
241	cattccaatc	accgctatca	ttgtatctgc	cctatgataa	gcgctgaaga	cccggacaga
301	ctacaacggt	atgcagaaa	acttaagaaa	agtgacatta	ccgacaagaa	catagcctct
361	aaggcggcag	acctgctgga	agtcattgtc	acaccagacg	cagagactcc	atctctgtgt
421	atgcacacag	acgccacgtg	taggtacttt	ggaagtgtag	cagtatacca	agatgtgtac
481	gcagtccatg	caccgacatc	aatctaccac	caggcgctta	aaggagttag	gacaatttac
541	tggataggct	ttgacacgac	cccttttatg	tacaaaaaca	tggcagggtc	ctaccctact

601	tacaacacga	actgggctga	cgagagagta	ttggaagcac	gtaacattgg	cctcggtaac
661	tcagatcttc	aggagagcag	gcttggaaaa	ctctcaatcc	ttaggaagaa	gaggctccaa
721	cctactaata	agatcatatt	ctcggttggt	tcaacaatct	acacagaaga	tagatcactg
781	ttacgtagct	ggcatcttcc	aaacgtgttc	cacttgaaag	gaaagtctaa	cttcacaggt
841	agatgtggga	ccattgtcag	ctgtgaaggg	tacgtcatca	aaaagataac	gatcagccca
901	ggactatacg	gtaaagttga	gaacttggcg	tccacaatgc	atcgcgaggg	tttcttgagt
961	tgcaaagtca	cagatacgtc	gcgcggcgag	agggtttctt	ttgctgtgtg	tacgtatgta
1021	ccagccacac	tttgcgatca	gatgacaggg	attctggcaa	ctgacgttag	tgtggatgac
1081	gcacaaaaac	tattgggttg	gctcaaccaa	aggattgtcg	tcaatggtag	gacgcaaaga
1141	aatactaaca	caatgcagaa	ctatctatta	ccagtggtcg	cccaggcggt	ttccagggtg
1201	gcgcgtgaac	atcgtgccga	cttggacgac	gagaaagaac	taggggtgcg	ggagcgcact
1261	cttactatgg	gctgctgctg	ggctttcaag	accagaaaaa	tcacatccat	ctacaagaag
1321	cctggtacgc	aaacaattaa	gaaagtacct	gccgtctttg	actcattttgt	gattccacgc
1381	cttaccagcc	acgggctcga	tatgggcttc	cgccgtaggc	tcaagctgct	gcttgaacca
1441	actgtcaaac	ccgcaccggc	tattacaatg	gccgatgtgg	agcatctgcg	tggcttacag
1501	caagaagctg	aagaagtggc	tgcagcggaa	gagatcagag	aagccctgcc	acccttgctc
1561	cctgaaatag	aaaaagagac	cgtagaggca	gaagtagacc	tcattatgca	agaggcagga
1621	gcaggtagcg	tggagacacc	acgaggacac	atcaggggtga	caagttaccc	aggcgaagag
1681	aagattgggt	cttacgctat	actttcaccc	caggcggtat	tgaatagtga	aaaactggcg
1741	tgtatccacc	cattggcgga	acaagtactg	gtaatgactc	acaaaggtag	ggcagggaga
1801	tacaaagtcg	agccatacca	cggtaaggtc	attgtaccag	aagggaacggc	ggtccctggt
1861	caagacttcc	aggcattgag	tgagagcgct	acgatcgttt	tcaacgagag	ggagttcgta
1921	aacagatacc	tgcaccacat	cgcaatcaac	ggaggagcgc	taaacactga	cgaagagtac
1981	tataagactg	taaagactca	ggacacagac	tcagaatacg	tcttcgatat	tgacgcacga
2041	aagtgtgtta	agcgagaaga	cgcaggtccc	ttgtgcctaa	ccggtgatct	ggtagatcca
2101	ccatttccacg	agtttgcgta	cgagagtctc	aagacacgac	cagcagcacc	tcacaaagtc
2161	ccaaccatcg	gagtctatgg	agtgccaggt	tcaggtaaat	ctggaatcat	caaaagcgct
2221	gtgactaaga	aagatctggt	tgtgagtgcg	aagaaggaaa	actgcgcaga	aatcatcagg
2281	gatgtaagga	ggatgagacg	tatggatggt	gctgctagga	ctgtcgattc	agtgtctcta
2341	aatgggggtta	agcaccocgt	taacactctg	tacattgatg	aggcatttgc	ctgccatgca
2401	gggacgctgc	tggcactgat	tgccatcgtc	aaacctaaaga	aagtgggtatt	gtgcggggac
2461	ccaaaacaat	gcggcttctt	taacatgatg	tgcctgaaag	tacattttta	ccatgacata
2521	tgcactgaag	tgtaccataa	aagcatctct	aggaggtgca	cacagactgt	aaccgccatc
2581	gtctccacgc	tcttctacga	caagcgaatg	aagacggtta	acccatgtgc	tgataaaatc
2641	atcatagata	ccacagggac	cacaaagccg	cacaaagatg	atctgattct	aacctgtttc
2701	agaggatggg	tgaacacagc	acagattgac	tacaaaaatc	acgaaatcat	gactgcggct
2761	gcacgcgaag	gacttacgcg	gaaaggcggt	tatgctgtca	ggtacaaagt	caacgagaat
2821	ccactctact	cgcagacttc	tgagcacgtg	aacgtgttac	ttacacgcac	agaaaaacgc
2881	attgtctgga	agacgctagc	tggtgatccc	tggataaaga	cacttacagc	taaatatccc
2941	ggggattttca	cggcttcatt	ggacgactgg	cagcgcgaac	acgacgccat	tatggcacgc
3001	gttcttgata	agccgcagac	agctgatgtg	ttccagaata	aggtgaacgt	ctgctgggcg
3061	aaggcttttag	agccagtcct	ggccacggcc	aacatttgtc	tgacgagaca	gcagtgggag
3121	acgttgcacc	cattcaagca	tgacagagcg	tactcacctg	aatggcact	gaacttcttt
3181	tgcaccaggt	tctttggagt	agacctggac	agtggttat	tttccgctcc	taccgtcgca
3241	cttacttaca	gggatcagca	ctgggataac	tgcgcaggga	agaacatgta	tgggcttaat
3301	agagaggttag	caaaggagtt	gtcacggcga	tatccgtgca	tcacaaaagc	ggttgacaca
3361	ggcagggtag	ctgatataag	gaataatacc	atcaaggact	actctccaac	aattaatgtg
3421	gttccattaa	atcgccgggt	gccccactcg	ttgatcgttg	accacaaagg	acagggtaga
3481	actgatcaca	gcggattcct	atctaagatg	aagggcaaat	ctgtgttggt	gatcggcgat
3541	cctatcagca	ttccagggaa	gaaagtagag	tccatgggtc	cattgcccac	taataccatc

3601	aggtgtgatc	tcgatttggg	aatacctagc	catgtcggta	aatatgacat	tatctttgtc
3661	aatgttagga	ccccgtacag	gaaccatcac	taccaacagt	gcgaggatca	cgctatccac
3721	cacagcatgc	taacgtgtaa	ggctgtccac	cacctgaaca	ctggcggaac	atgtgtggct
3781	ataggggatg	ggcttgctga	tcgcgcaacc	gagaatatca	tcactgcggt	ggcacgctca
3841	tttaggttta	cccgtgtctg	tcagcctaag	aacactgccg	aaaatactga	ggttctcttc
3901	gtgtttcttcg	gcaaggacaa	cggcaaccac	acacatgacc	aggacagact	cgggtgtagt
3961	cttgacaaca	tctatcaagg	gtcaaccagg	tacgaggcag	ggagagctcc	agcgtaacaga
4021	gtgatcagag	gtgacattag	caagagcgct	gaccaagcta	tcgttaatgc	tgctaatagc
4081	aaaggtcaac	caggttcg	agtgtgcgg	gcactgtacc	gaaaatggcc	ggctgctttt
4141	gatagacagc	caatagctgt	cgggacggct	agacttgtga	agcacgaacc	gctcatcata
4201	catgctgtag	gacccaattt	ttctaagatg	ccggaaccgg	agggcgacct	taagctcgca
4261	gctgcctaca	tgagcatagc	gtccatcgtc	aacgctgagc	ggattacaaa	aatatcagta
4321	ccgctactgt	caaccggcat	ctattctgg	ggcaaagatc	gagtgatgca	atcattgcat
4381	cacctgttca	ctgctttcga	cactacggat	gccgatgtca	ccatatattg	cttggataaa
4441	caatgggaga	ccaggataat	cgaggccatt	caccgcaaag	aaagcgtcga	aattctggat
4501	gatgacaagc	cagtagacat	tgacttgg	agggctccacc	caaacagctc	tttggcaggg
4561	agaccagggt	actccgtcaa	tgagggcaag	ttgtattcat	acctggaagg	tacacgattc
4621	catcagaccg	ccaaggacat	tgccgaaatc	catgcaatgt	ggcccaacaa	atctgaggct
4681	aatgagcaga	tttgcttgta	cactctggg	gagagtatgt	ccagcatccg	ctccaaatgc
4741	ccagtagagg	agtcagaggc	gtctgtcca	cctcacacac	ttccatgcct	gtgtaattac
4801	gctatgacgg	ctgagcgct	atacaggttg	cgctctgcga	agaaagaaca	gttcgccgta
4861	tgctcatcat	tcctgttgcc	gaagtacagg	atcacaggcg	tgcagaagct	acagtgcagc
4921	aaaccagtcc	tggttttcagg	cgctgtacca	ccggctgtac	accccaggaa	gtacgcggaa
4981	ataattctag	aaacgccacc	accgccagca	acgacaaccg	taatattgtga	accaactgtg
5041	ccagaacgta	taccaggtcc	ggtgatttct	agagcaccaa	gtgcggaate	actgctatcg
5101	cttggcgggc	tctcgttctc	tagctctgcc	acacgctcgt	caaccgcctg	gagcgactat
5161	gacaggcggt	ttgtgggttac	agctgatgtg	catcaagcga	acacgtctac	gtggagcatc
5221	cctagtgtc	ctggcttgga	cgctcagctg	ccttctgacg	tcactgatte	ccactggagt
5281	attccaagtg	catcaggctt	tgaagtgaga	acaccatctg	tacaggacct	aactgcggag
5341	tgtgcgaagc	ctcgtggact	ggccgaaata	atgcaagact	tcaatactgc	tcctttccag
5401	tttctttcgg	actacagacc	agtaccggca	ccacggagac	gccccatccc	atcacctaga
5461	tcgacggctt	ccgcacctcc	agttccaaag	ccacgcagga	ctaagtacca	acaaccacca
5521	ggagtgccta	gagcgatctc	agaagcggag	ttggacgagt	acatccgtca	acactccaac
5581	tgacggtatg	aagcgggagc	gtatattttc	tcctcggaaa	caggccaagg	tcaccttcaa
5641	cagaaatcag	tacgtcaatg	taaactacaa	gaacctatat	tggatcgggc	cgtccatgag
5701	aagtattacg	ccccgcgcct	cgatctcgaa	agagagaaaa	tgttacagaa	gaaactgcaa
5761	ttatgcgcct	ctgaaggaaa	tagaagcagg	tatcaatcac	gaaaagtaga	aaatatgaaa
5821	gcaattacag	cggagcgact	catttctgga	ttgggacacat	atctatcatc	agaagtgaat
5881	cctgtcgagt	gttacagagt	caattatcct	gtaccaatct	actcgtcaac	ggtaatatac
5941	aggttttacat	ctgcagaggt	cgcggttaaa	acgtgcaact	tagttatcca	agagaattac
6001	cctacagtag	ccagttattg	tataacagat	gaatacagatg	cgtatcttga	catgggtggac
6061	ggcgcatcgt	gctgtctaga	tacagccact	ttttgtccgg	ctaaactgag	aagctacca
6121	aagaagcata	gctatttgca	gccagagata	agatcagccg	tcccatcgcc	tatacagaat
6181	acattacaaa	atgtattggc	tgacgctact	aaaaggaatt	gcaacgttac	ccaaatgcga
6241	gaattacctg	tcttagatc	ggcggcattt	aatgttgatt	gtttcaagaa	atagcatgc
6301	aatgatgagt	actgggatac	ctttcgcgat	aaccctatc	ggctaactac	agagaacgtt
6361	acgcaatatg	tgacaaagct	gaaagggccg	aaagcagcag	cattgtttgc	gaatactcat
6421	aatctaaaac	cgttgcagga	gataccaatg	gatcaattcg	tcattggatct	aaagagagat
6481	gtcaaagtta	ctcccggcac	gaaacataca	gaggagcggc	ctaaggtgca	ggttattcag
6541	gctgcagatc	cccttgctac	cgcttacctt	tgcgggatcc	atcggaatt	agtccgtaga

6601	ctgaatgcgg	tgcttctgcc	aaatatccat	actctcttcg	acatgtcagc	ggaagatttt
6661	gatgcgatta	ttgctgaaca	tttccaccac	ggcgacccag	tattggaaac	ggacatcgcg
6721	tcgtttgata	aaagcgaaga	cgacgctatc	gccatttcgg	cgttgatgat	ccttgaggac
6781	ttaggtgtcg	accaaccgct	cttagatttg	atagaggcgg	cgttcggcaa	tatcacatct
6841	gtgcacctac	ctacaggaac	gaggtttaaa	tttggtgcca	tgatgaaatc	cggtatgttc
6901	ttaacgctgt	ttgtcaacac	actagtcaat	atcatgattg	ctagcagagt	actacgtgaa
6961	cggttaacca	cgtcagcgtg	cgcggcctct	atcggcgacg	ataacatagt	gcatggtgtc
7021	gtctccgaca	ccttgatggc	ggagagatgc	gccacttggc	tgaacatgga	agtaaaaatt
7081	attgatgcag	ttattggtat	caaagcacc	tacttctgtg	ggggatttat	cctggtggac
7141	cagataacag	gcacagcctg	cagagtcgca	gaccctctaa	aaaggctttt	taagcttgga
7201	aaaccattgc	cagtcgatga	tacccaagac	tgcgaccgcc	gccgggcact	gcatgatgaa
7261	gcaatgcgat	ggaacagaat	tgggaattacg	gacgagttag	tgaaggccgt	agaatccaga
7321	tacgagatca	tactggcagg	cctgatcatc	acgtctctgt	ccacgttagc	cgaaagcggt
7381	aagaacttca	agagcataag	agggagccca	atcaccctct	acggctgacc	taaatagggtg
7441	acgtagtaga	cacgcacctc	cccaccggca	gaatgtttcc	ataccctcag	ctgaactttc
7501	caccagttta	ccctacaaat	ccgatggctt	accgagatcc	aaaccctcct	aggcgccgct
7561	ggaggccggt	tcggcccccgc	ctggctgctc	aaatcgaaga	tcttaggagg	tcgatagtca
7621	acttgacttt	caaacaacga	tcacctaatc	cgccgccagg	tccaccgcca	aagaagaaga
7681	agagtgtcc	taagccaaaa	cctactcagc	ctaaaaagaa	gaagcagcaa	gccaagagga
7741	cgaaacgcaa	gcctaaacca	gggaaacgac	aacgtatgtg	tatgaagttg	gagtcggaca
7801	agacatttcc	gatcatgctg	aacggccaag	tgaatggata	tgcctgctgt	gtcggaggaa
7861	ggctgatgaa	accactccac	gttgaaggaa	aaattgataa	tgagcaatta	gcgcccggtga
7921	aattgaagaa	ggctagcatg	tacgacttgg	agtacggcga	cgttccccag	aacatgaaat
7981	cagacacgct	gcagtacacc	agcgacaaac	caccgggctt	ctacaactgg	caccacggcg
8041	cagtcacgta	tgagaatggg	agatttaccg	taccgagagg	agtgggcggg	aaaggcgaca
8101	gcggaagacc	gatcctggac	aacagaggca	gagttgtggc	tattgttcta	ggaggtgcaa
8161	atgagggcac	gcgtacggcg	ctttcagtg	tcacttggaa	ccagaaagg	gtgaccatta
8221	gggatacccc	cgaagggttct	gaaccgtgg	cactagttac	agcgctatgc	gtgctttcga
8281	atgtcacggt	cccattgcgac	aaaccacccg	tgtgctattc	actgacgcca	gaacgaacac
8341	tcgacgtgct	cgaagagaac	gtcgacaatc	caaattacga	cacgctgctg	gagaacgtct
8401	tgaatgtcc	atcacgccgg	cccaaacgaa	gcattaccga	tgacttcaca	ctgaccagtc
8461	cctacctggg	gttctgccc	tattgcagac	actcaacgcc	gtgtttcagc	ccaataaaaa
8521	ttgagaacgt	gtgggacgaa	tctgatgatg	gatcgattag	aatccaggct	tcggcacaa
8581	tcggctacaa	tcaggcaggc	actgcggatg	tcaccaaatt	ccgttacatg	tctttcgacc
8641	acgaccatga	catcaaggaa	gacagtatgg	agaaaatagc	tatcagcaca	tctggaccct
8701	gccgtcgtct	tggccacaaa	gggtacttcc	tgtagctca	atgtcctcca	ggtgacagtg
8761	taaccgtcag	tatcacgagc	ggagcatctg	agaattcatg	caccgtggag	aaaaagatca
8821	ggaggaagtt	tgtcggtaga	gaggagtact	tgttcccacc	cgtccatgga	aagctggtaa
8881	agtgccacgt	ttacgatcac	ttgaaggaga	cgtctgccgg	gtacataacc	atgcacaggc
8941	caggcccaca	cgcgtataag	tcctatctgg	aggaagcgtc	aggcgaagtg	tacattaaac
9001	caccttctgg	caagaacgtc	acctacgaat	gtaagtgtgg	cgactacagc	acaggtatcg
9061	tgagcacgcg	aacgaagatg	aacggctgca	ctaaagcaaa	acagtgcatt	gcctacaaga
9121	gcgaccaaac	gaaatgggtc	ttcaactcgc	cggatcttat	taggcacaca	gaccactcag
9181	tgcaaggtaa	attgcacatt	ccattccgct	tgacaccgac	agtctgccc	gttccggttag
9241	ctcacacgcc	tacagtcacg	aagtggttca	aaggcatcac	cctccacctg	actgcaatgc
9301	gaccaacatt	gctgacaacg	agaaaattgg	ggctgcgagc	agacgcaaca	gcagaatgga
9361	ttacagggtc	tacatccagg	aatttttctg	tggggcgaga	agggctggag	tacgtatggg
9421	gtaaccatga	accagtcaga	gtctggggccc	aggagtcggc	accaggcgac	ccacatggat
9481	ggccgcgatga	gatcatcatc	cactattatc	atcggcatcc	agtctacact	gtcattgtgc
9541	tgtgtggtgt	cgctcttgct	atcctggtag	gcactgcate	atcagcagct	tgcacgcca

```

9601 aagcaagaag agactgcctg acgccatacg cgcttgcacc gaacgcaacg gtacccacag
9661 cattagcggg tttgtgctgc attcggccaa ccaacgctga aacatttgga gaaactttga
9721 accatctgtg gtttaacaac caaccgtttc tctgggcaca gttgtgcatt cctctggcag
9781 cgcttggtat tctgttccgc tgcttttcat gctgcattgcc ttttttattg gttgcaggcg
9841 tctgcctggg gaaggtagac gccttcgaac atgcgaccac tgtgccaaat gttccgggga
9901 tcccgtataa ggcgttggtc gaacgcgcag gttacgcgcc acttaacctg gagatcacgg
9961 tcgtctcatc ggaattaaca ctttcaacta acaaggagta cgtgacctgc aaattccaca
10021 cagtcattcc ttcaccacaa gttaaatgct gcgggtccct cgagtgcagg gcatcctcaa
10081 aggcggaata cacatgccgc gtttttggcg gtgtgtaccc tttcatgtgg ggaggcgcac
10141 aatgcttctg tgacagttag aacacacaac tgagttaggc gtacgtcgag ttcgctccag
10201 actgcactat agatcacgca gtcgcactaa aagttcacac agctgctctg aaagtcggcc
10261 tgcgtatagt atacggcaac accaccgcgc acctggatac gtttgtcaat ggcgtcacgc
10321 caggttcctc acgggacctg aaggtcatag cagggccgat atcagccgct ttttcacctt
10381 ttgaccataa ggtcgtcatc agaaaggggc ttgtttacaa ctacgacttc cctgagtatg
10441 gagctatgaa accaggagcg ttcggcgata ttcaagcate ctcgcttgat gctacagaca
10501 tagtagcccg cactgacata cggctgctga agccttctgt caagaacatc cacgtccctt
10561 acaccaagc agtatcaggg tatgaaatgt ggaagaacaa ctcaggacga cccctgcaag
10621 aaacagcacc atttggatgt aaaattgaag tggagcctct gcgagcgtct aactgtgctt
10681 acgggcacat ccctatctcg attgacatcc ctgatgcagc ttttgtgaga tcatcagaat
10741 caccaacaat tttagaagtt agctgcacag tagcagactg catttattct gcagactttg
10801 gtggttctct aacattacag tacaaagctg acaggaggag acattgtcca gttcactccc
10861 actccacgac agctgttttg aaggaagcga ccacacatgt gactgccgta ggcagcataa
10921 cactacattt tagcacatcg agcccacaag caaattttat agtttcgcta tgcggcaaga
10981 agtccacctg caatgctgaa tgtaaaccac cggccgacca cataattgga gaaccacata
11041 aagtcgacca agaattccag gcggcagttt ccaaaacatc ttggaactgg ctgcttgcac
11101 tgtttggggg agcatcatcc ctcatgttg taggacttat agtgttgggtc tgcagctcta
11161 tgcttataaa cacacgtaga tgactgagcg cggacactga catagcggta aaactcgatg
11221 tacttccgag gaagcgtggt gcataatgcc acgcgccgct tgacactaaa actcgatgta
11281 tttccgagga agcacagtgc ataatgctgt gcagtgtcac attaatcgta tatcacacta
11341 catattaaca acactatatc acttttatga gactcactat gggctctctaa tatacactac
11401 acatatttta cttaaaaaaca ctatacacac tttataaatt cttttataat ttttctttt
11461 tttttatattt gtttttaaaa tttc

```

//



SEQUENCE LISTING - pcDWXH-7 - SEQ ID NO 2

<210> SEQ ID NO 2  
 <211> LENGTH: 4150  
 <212> TYPE: DNA  
 <213> ORGANISM: WESTERN EQUINE ENCEPHALITIS VIRUS  
 <214> FEATURES:

Location/Qualifiers

source 1..4150  
 /organism="Western equine encephalomyelitis virus"  
 /strain="71V-1658"  
 vector sequence 1-9  
 5' SacI primer 9-20  
 3' end of NS4 gene 16-114  
 intragenic region 115-159  
 CDS 160-3869  
 /note="encodes capsid, E3, E2, 6K and E1 proteins"  
 /codon\_start=160  
 /product="structural polyprotein"

/translation="MFPYPQLNFPPVYPTNPMA YRDPNP PRRRWRPFRPPLAAQIEDL  
 RRSIVNLTFKQRSPNPPPGPPP KKKKSAPKPKPTQPKKKKQQA KRTKRKPKPGKRQRM  
 CMKLESDKTFPIMLNGQVNGYACVVGGRLMKPLHVEGKIDNEQLAAVKLKKASMYDLE  
 YGDVPQNMKSDTLQYTS DKPPGFYNWHHGAVQYENGRFTVPRGVGGKGDSGRPI LDNR  
 GRVVAIVLGGANEGTRTALS SVVTWNQKGV TIRDTPEGSEPW SLVTALCVLSNVTFPCD  
 KPPVCYSLTPERTLDVLEENVDNPNYDTLLENVLKCP SRRPKRSITDDFTLTSPYLG F  
 CPYCRHSTPCFSP IKIENVWDESDDGSIRIQVSAQFGYNQAGTADVTKFRYMSFDH DH  
 DIKEDSMEKIAISTSGPCRRLGHKGYFLLAQC PPGDSVTVSITSGASENSCTVEKKIR  
 RKFVGREEYLFPPVHGKLVKCHVYDHLKETSAGYITMHRPGPHAYKSYLEEASGEVYI  
 KPPSGKNVTYECKCGDYSTGIVSTRTKMNGCTKAKQCIAYKSDQTKWVFNSPD LIRHT  
 DHSVQGKLH I PFRLTPTVCPVPLAHTPTVTKWFKGITLHLTAMRPTLLTTRKLGLRAD  
 ATAEWITGSTSRNFSV GRELEYVWGNHEPVRVWAQESAPGDPHGWPHEII IHYYHRH  
 PVYTVIVLCGVALAILVGTASSAACIAKARRDCLTPYALAPNATVPTALAVLCCIRPT  
 NAETFGETLNHLWFNNQPFLWAQLCIPLAALVILFRCFSCCMPFLLVAGVCLGKVDAF  
 EHATTVPNVPGIPYKALVERAGYAPLNLEITVVSSELTPSTNKEYVTCKFHTVIPSPQ

10023649-122101

VKCCGSLECKASSKADYTCRVFGGVYPFMWGGAQCFCDSSENTQLSEAYVEFAPDCTID  
 HAVALKVHTAALKVGLRIVYGNTTAHLDTFVNGVTPGSSRDLKVIAGPISAAFPSFDH  
 KVVIRKGLVYNYDFPEYGAMKPGAFGDIQASSLDATDIVARTDIRLLKPSVKNIHVPY  
 TQAVSGYEMWKNNNGRPLQETAPFGCKIEVEPLRASNCAYGHIPI SIDIPDAAFVRSS  
 ESPTILEVSCTVADCIYSADFGGSLTLQYKADREGHCPVHSHSTTAVLKEATTHVTAV  
 GSITLHFSTSSPQANFIVSLCGKKSTCNAECKPPADHIIGEPHKVDQEFQAAVSKTSW  
 NWLLALFGGASSLIVVGLIVLVCSSMLINTRR"

3'UTR

3870-4150

BASE	COUNT	ORIGIN
1	ggccctctag	agctcatact
61	agcgtaaaga	acttcaagag
121	taggtgacgt	agtagacacg
181	actttccacc	agtttaccct
241	gccgctggag	gccgttttcgg
301	tagtcaactt	gacttttcaaa
361	agaagaagag	tgctcctaag
421	agaggacgaa	acgcaagcct
481	cggacaagac	atttccgatc
541	gaggaaggct	gatgaaacca
601	ccgtgaaatt	gaagaaggct
661	tgaaatcaga	cacgctgcag
721	acggcgcagt	ccagtatgag
781	gcgacagcgg	aagaccgatc
841	gtgcaaattga	gggcacgcgt
901	ccattagggg	tacccccgaa
961	tttcgaatgt	cacgttccca
1021	gaacactcga	cgtgctcgaa
1081	acgtcttgaa	atgtccatca
1141	ccagtcacct	cctgggggttc
1201	taaaaattga	gaacgtgtgg
1261	cacaattcgg	ctacaatcag
1321	tcgaccacga	ccatgacatc
1381	gaccctgccg	tcgtcttggc
1441	acagtgtaac	cgtcagtatc
1501	agatcaggag	gaagtttgtc
1561	tggtaaagtg	ccacgtttac
1621	acaggccagg	cccacacgcg
1681	ttaaaccacc	ttctggcaag
1741	gtatcgtgag	cacgcgaacg
1801	acaagagcga	ccaaacgaaa
1861	actcagtgca	aggtaaattg
1921	cgtttagctca	cacgcctaca

10023649.122101

1981 caatgcgacc aacatttgctg acaacgagaa aattggggct gcgagcagac gcaacagcag  
 2041 aatggattac agggctctaca tccaggaatt tttctgtggg gcgagaaggg ctggagtacg  
 2101 tatggggtaa ccatgaacca gtcagagtct gggcccagga gtcggcacca ggcgaccac  
 2161 atggatggcc gcatgagatc atcatccact attatcatcg gcatccagtc tacactgtca  
 2221 ttgtgctgtg tgggtgctgct cttgctatcc tggtaggcac tgcacatca gcagcttgca  
 2281 tcgccc aaagc aagaagagac tgccctgacgc catacgcgct tgcaccgaac gcaacgggtac  
 2341 ccacagcatt agcgggttttg tgctgcattc ggccaaccaa cgctgaaaca tttggagaaa  
 2401 ctttgaacca tctgtgggtt aacaaccaac cgtttctctg ggcacagttg tgcattcctc  
 2461 tggcagcgct tggtattctg ttcgctgct tttcatgctg catgcctttt ttattgggtg  
 2521 caggcgtctg cctggggaag gtagacgct tcgaacatgc gaccactgtg ccaaattgtt  
 2581 cggggatccc gtataaggcg ttgggtcgaac gcgcagggtta cgcgccactt aacctggaga  
 2641 tcacggctcg ctcacggaa ttaacacctt caactaacia ggagtacgtg acctgcaaat  
 2701 tccacacagt cattccttca ccacaagtta aatgctgcgg gtccctcgag tgcaaggcat  
 2761 cctcaaaggc ggattacaca tgccgcgttt ttggcggtgt gtaccctttc atgtggggag  
 2821 gcgcacaatg cttctgtgac agtgagaaca cacaactgag tgaggcgtag gtcgagttcg  
 2881 ctccagactg cactatagat cacgcagtcg cactaaaagt tcacacagct gctctgaaag  
 2941 tcggcctgcg tatagtatac ggcaacacca ccgcgcacct ggatacgttt gtcaatggcg  
 3001 tcacgccagg ttcctcacgg gacctgaagg tcatagcagg gccgatatca gccgcttttt  
 3061 caccctttga ccataaggct gtcacagaa aggggcttgt ttacaactac gacttccttg  
 3121 agtatggagc tatgaaacca ggagcgctcg gcgatattca agcatcctcg cttgatgcta  
 3181 cagacatagt agcccgact gacatacggc tgctgaagcc ttctgtcaag aacatccacg  
 3241 tcccctacac ccaagcagta tcagggtatg aaatgtggaa gaacaactca ggacgacccc  
 3301 tgcaagaaac agcaccattt ggatgtaaaa ttgaagtgga gcctctgcca gcgtctaact  
 3361 gtgcttacgg gcacatccct atctcgattg acatccctga tgcagctttt gtgagatcat  
 3421 cagaatcacc aacaattttta gaagttagct gcacagtagc agactgcatt tattctgcag  
 3481 actttgggtg ttctctaaca ttacagtaca aagctgacag ggagggacat tgtccagttc  
 3541 actcccactc cacgacagct gttttgaagg aagcgaccac acatgtgact gccgtaggca  
 3601 gcataaact acatttttagc acatcgagcc cacaagcaaa ttttatagtt tcgctatgcg  
 3661 gcaagaagtc cacctgcaat gctgaatgta aaccaccggc cgaccacata attggagaac  
 3721 cacataaagt cgaccaagaa ttccaggcgg cagtttccaa aacatcttgg aactggctgc  
 3781 ttgcactggt tgggggagca tcatccctca ttgttgtagg acttatagtg ttggtctgca  
 3841 gctctatgct tataaacaca cgtagatgac tgagcgcgga cactgacata gcggtaaaac  
 3901 tcgatgtact tccgaggaag cgtggtgcat aatgccacgc gccgcttgac actaaaactc  
 3961 gatgtatttc cgaggaagca cagtgcataa tgctgtgcag tgtcacatta atcgtatatc  
 4021 aactacata ttaacaacac tatatcactt ttatgagact cactatgggt ctctaataa  
 4081 cactacacat attttactta aaaacactat acacacttta taaattctct cataatttca  
 4141 ctttaggttt  
 //

SEQUENCE LISTING - pVHX-6 - SEQ ID NO 3

<200> SEQUENCE CHARACTERISTICS  
 <210> SEQ ID NO 3  
 <211> LENGTH: 1260  
 <212> TYPE: DNA  
 <213> ORGANISM: EXAMPLE OF A CMV PROMOTER COUPLED TO THE pcDW-XH7  
 SEQUENCE

<220> FEATURE:  
 1-302 pVAX vector sequence  
 302-401 CMV promoter  
 411 CMV putative transcriptional start site  
 444-463 T7 promoter  
 464-509 pVAX multicloning region  
 510 start of pcDW-HX45 sequence

10023649.122101

1	cattgacgtc	aatgggtgga	gtattttacgg	taaactgccc	acttggcagt	acatcaagtg
61	tatcatatgc	caagtacgcc	ccctattgac	gtcaatgacg	gtaaatggcc	cgcctggcat
121	tatgcccagt	acatgacctt	atgggacttt	cctacttggc	agtacatcta	cgtattagtc
181	atcgctatta	ccatgggtgat	gcgggttttg	cagtacatca	atggggcgtgg	atagcggttt
241	gactcacggg	gattttccaag	tctccacccc	attgacgtca	atggggagttt	gttttggcac
301	caaaatcaac	gggactttcc	aaaatgtcgt	aacaactccg	ccccattgac	gcaaatgggc
361	ggtaggcgtg	tacgggtggga	ggcttatata	agcagagctc	tctggctaac	tagagaaccc
421	actgcttact	ggcttatcga	aattaatacg	actcactata	gggagaccca	agctggctag
481	cgttttaaact	taagcttggg	accgagctca	tactggcagg	cctgatcatc	acgtctctgt
541	ccacgtttagc	cgaaagcggt	aagaacttca	agagcataag	agggagccca	atcacccctct
601	acggctgacc	taaataggtg	acgtagtaga	cacgcaccta	cccaccgccca	gaatgtttcc
661	ataccctcag	ctgaactttc	caccagttta	ccctacaaat	ccgatggcct	accgagatcc
721	aaaccctcct	aggcgccgct	ggaggccggt	tcgggccccc	ctggctgctc	aaatcgaaga
781	tcttaggagg	tcgatagtca	acttgacttt	caaacaacga	tcacctaata	cgccgccagg
841	tccaccgcca	aagaagaaga	agagtgtccc	taagccaaaa	cctactcagc	ctaaaaagaa
901	gaagcagcaa	gccaagagga	cgaaacgcaa	gcctaaacca	gggaaacgac	aacgtatgtg
961	tatgaagttg	gagtcggaca	agacatttcc	gatcatgctg	aacggccaag	tgaatggata
1021	tgcttgcgtt	gtcggaggaa	ggctgatgaa	accactccac	ggtgaaggaa	aaattgataa
1081	tgagcaatta	gcggccgtga	aattgaagaa	ggctagcatg	tacgacttgg	agtacggcga
1141	cgttccccag	aacatgaaat	cagacacgct	gcagtacacc	agcgacaaac	caccgggctt
1201	ctacaactgg	caccacggcg	cagtccagta	tgagaatggg	agatttaccg	taccgagagg

. . . . CONTINUOUS WITH pcDW-HX45 (pcDW-XH7) SEQUENCE . . . . .

4177	gaatgtaaac	caccggccga	ccacataatt	ggagaaccac	ataaagtcga	ccaagaattc
4237	caggcggcag	tttccaaaac	atcttgggaac	tggtctgcttg	cactgttttg	gggagcatca
4297	tccctcattg	ttgtaggact	tatagtgttg	gtctgcagct	ctatgcttat	aaacacacgt
4357	agatgactga	gcgcggacac	tgacatagcg	gtaaaactcg	atgtacttcc	gaggaagcgt
4417	ggtgcataat	gccacgcgaa	gcttgacact	aaaactcgat	gtattttccga	ggaagcacag
4477	tgcataatgc	tgtgcagtgt	cacattaatc	gtatatcaca	ctacatatta	acaacactat
4537	atcactttta	tgagactcac	tatgggtctc	taatatacac	tacacatatt	ttacttaaaa
4597	acactataca	cactttataa	attcttttat	aattttttctt	ttgcttttaga	gcacactggc

4657 ggccggttact agtggatccg agctctagag ggcc

4177-4643 pcDW-HX45

4644-4790 vector sequence //

10023649.123101